TECOLOTE RESEARCH, INC.

Bridging Engineering and Economics

How Many Iterations Are Enough?

Alfred Smith

2008 NASA Cost Symposium

Cost Estimating: "It's a Risky Business", Portland, OR

August 2008

■ Los Angeles ■ Washington, D.C. ■ Boston ■ Chantilly ■ Huntsville ■ Dayton ■ Santa Barbara ■ Albuquerque ■ Colorado Springs ■ Ft. Meade ■ Ft. Monmouth ■ Ogden ■ Silver Spring ■ Patuxent River ■ Pensacola ■ Washington Navy Yard ■ Goddard Space Flight Center ■ Cleveland ■ Denver ■ Johnson Space Center ■ Dahlgren ■ Montgomery ■ New Orleans ■ Oklahoma City ■ Tampa ■ Tacoma ■ Warner Robins ALC ■ Vandenberg AFB

Outline

- **Setting the Stage**
  - Describe a cost uncertainty simulation model
- **How to Test for Convergence**
  - Analytic test for convergence
  - Test for convergence using simulation data
  - Propose a simple, repeatable, tool independent approach
- **Applying the Approach to Several Models**
  - Look for patterns
  - Identify model characteristics that influence the iterations required
- **Concluding Comments**

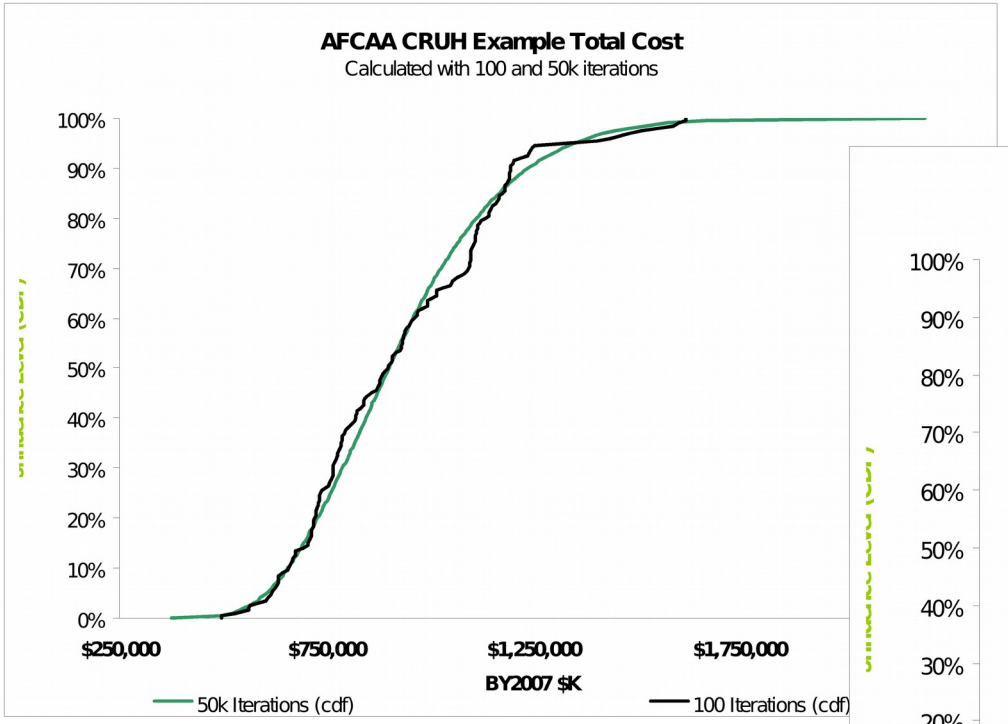
A Sample "Inputs Simulation" Model: AFCAA CRUH Example File

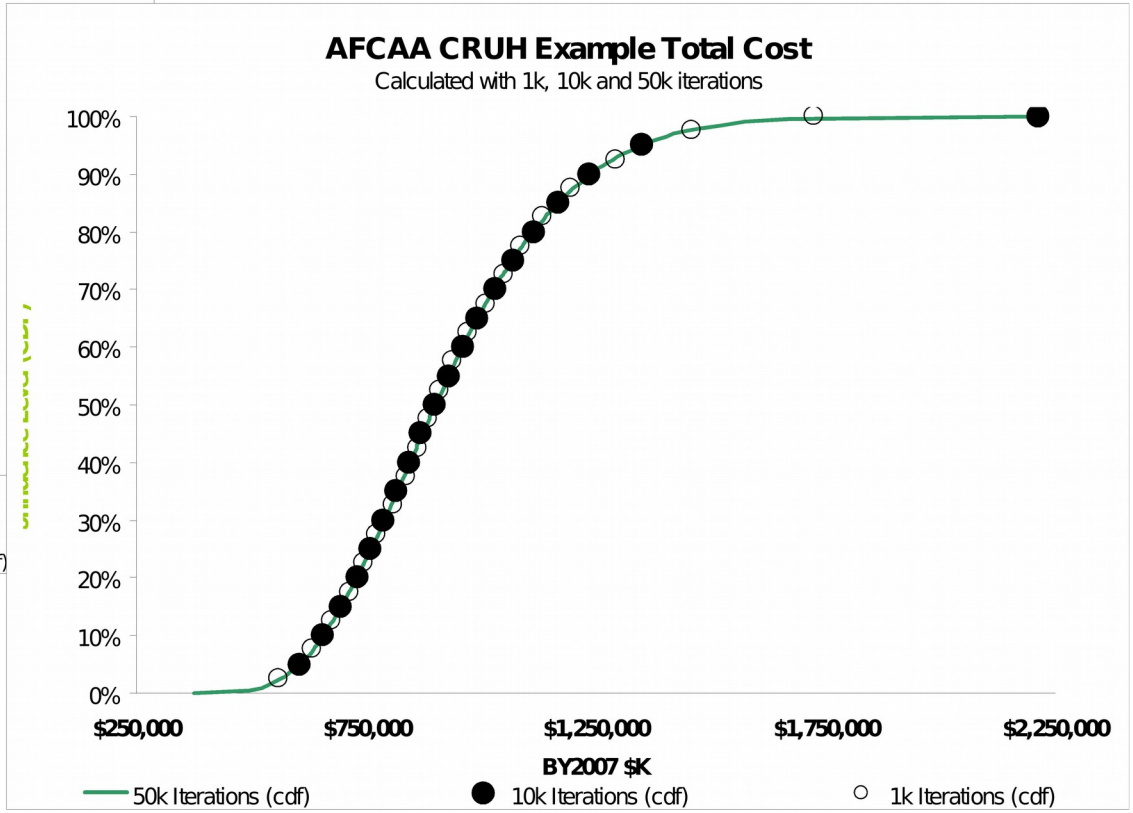
- Missile System
  - Sys Dev and Demo
    - Air Vehicle
      - Design & Development
      - Prototypes
      - Software
    - Sys Engineering/Program Management
    - System Test and Evaluation
    - Training
    - Data
    - Support Equipment
  - Production Phase
    - Air Vehicle
      - Payload
      - Propulsion
      - Airframe
      - Guidance and Control
      - Integration, Assembly, Test and Checkout
    - Engineering Changes
    - Sys Engineering/Program Management
    - System Test and Evaluation
    - Training
    - Data
    - Peculiar Support Equipment
    - Common Support Equipment
    - Initial Spares and Repair Parts

- Total For Discrete Uncertainties
  - CSE Item #1 60%
  - CSE Item #2 60%
  - CSE Item #3 10%
  - CSE Item #4 30%
  - CSE Item #5 10%
  - CSE Item #6 50%
  - CSE Item #7 50%
  - CSE Item #8 10%
  - CSE Item #9 40%
  - CSE Item #10 50%

- **21 WBS elements (lowest level), 38 input variables**
- **Most of the common estimating methods are represented**
  - Linear, loglinear, triad, factor, build-up, third party tools, throughputs
  - Date driven methods (uncertainty on duration)
- **Normal, lognormal, triangular, uniform uncertainty distributions**
  - Functional and applied correlation
  - Includes 10 discrete (Bernoulli) distributions
- **Modeled using @Risk, CB, ACEIT**

S-Curves Based On Different Iterations





- 100 iterations clearly not enough
- 1k iterations almost matches the 50k run
- No visual evidence that 10k any different from 50k iteration result

TECOLOTE RESEARCH, INC.

Bridging Engineering and Economics

Convergence

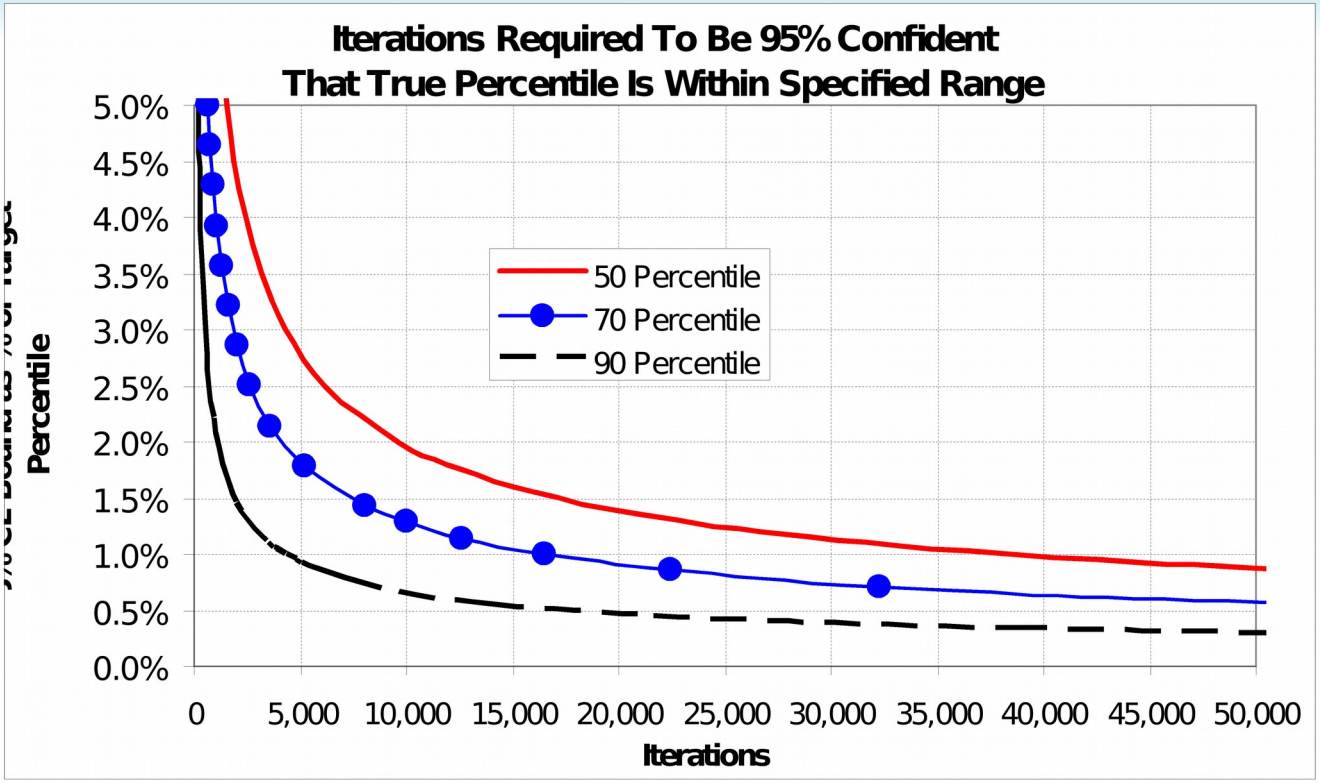
Analytical Solution

$$m = p(1- p)\left[\frac{c}{\Delta p}\right]^2$$

- **Where:**
  - m = number of iterations
  - p = the percentile of interest
  - c = inverse of the standard normal cumulative distribution
    - For 95% confidence, in Excel use Normsinv(0.975)
  - Δp = the percentile range of interest (for instance, use 0.05 if interested in +/- 5 percentile)
- **Independent of distribution shape**

Source: M Granger Morgan and Max Henrion, UNCERTAINTY, A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis, pp 202

Analytical Solution



Observations

- More iterations required to converge on the 50 percentile than 90 percentile
  - Morgan & Henrion pp 202 describe the 50 percentile as “the least precise estimated percentile”
- Need 5 to 35k iterations to have error less than 1%
- Will Latin Hypercube sampling improve on this result?

Test for Sufficient Iterations From Simulation Data

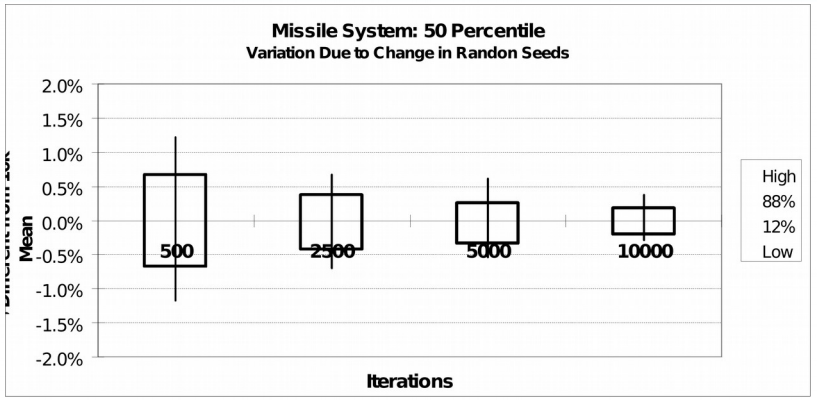
- **Goal:** create a simple way to determine sufficient number of iterations to obtain "accurate" results using the simulation data
- **Several potential metrics of interest:**
  - Mean, standard deviation, coefficient of variation
  - Correlation coefficient
  - Target percentile
  - Other? All?

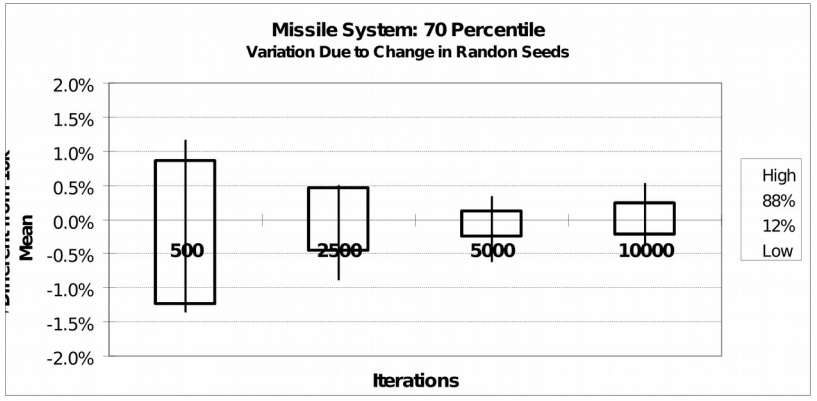
- **Selected:** target percentile for the WBS element(s) of interest
  - Selected because this is the result that tends to be the basis for budget recommendations
  - 50%, **70%**, 90% used in this study, but the one your decision maker needs might be a better choice

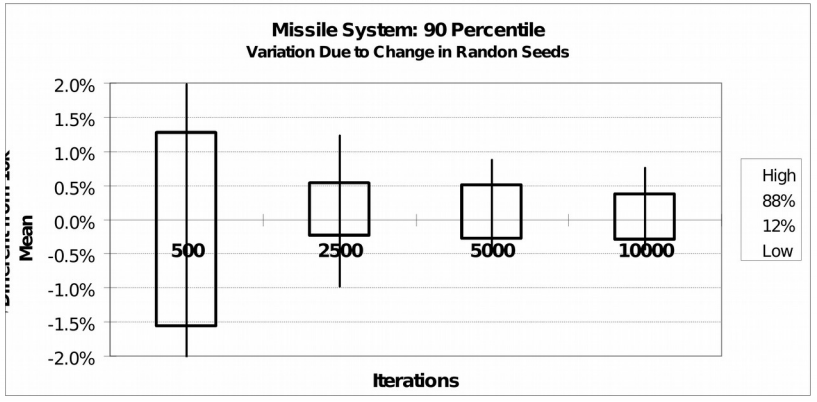
Several Issues to Resolve:

- **How do we know the "right answer"**
  - Comparing a complex cost model simulation result to an analytic solution is not feasible
  - Literature identifies 10k iterations as the benchmark for "sufficient"
    - Morgan MG, Henrion M (1990) Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis
    - Garvey P (2000) Probability Methods for Cost Uncertainty Analysis: A Systems Engineering Perspective
- **How to gather the data?**
  - Use Latin Hypercube sampling rather than Monte Carlo
  - Is it necessary to change the random seeds on each run?
  - Is it necessary to perform separate runs, or is the data from a single 10k run sufficient?
- **How to present results? Options include:**
  - Plot multiple statistics for a specific result
  - Plot single statistic for multiple results
  - Plot x iteration result as a % difference from the "correct" result
  - **Selected:** Plot x iteration result as the absolute % difference to the "correct" result

Is it necessary to change random seed when checking for convergence?







- 25 identical CRUH files, but with a different set of random seeds
- All 25 files run at 500, 2500, 5000 and 10000 iterations
- 50, 70 and 90 percentile results at the Total level each compared to the average of the 10k result across all 25 files
- **Observation**: random seed selection generally has less than +/- 0.5% impact on most results
- **Conclusion:** No need to change random seed to check for convergence

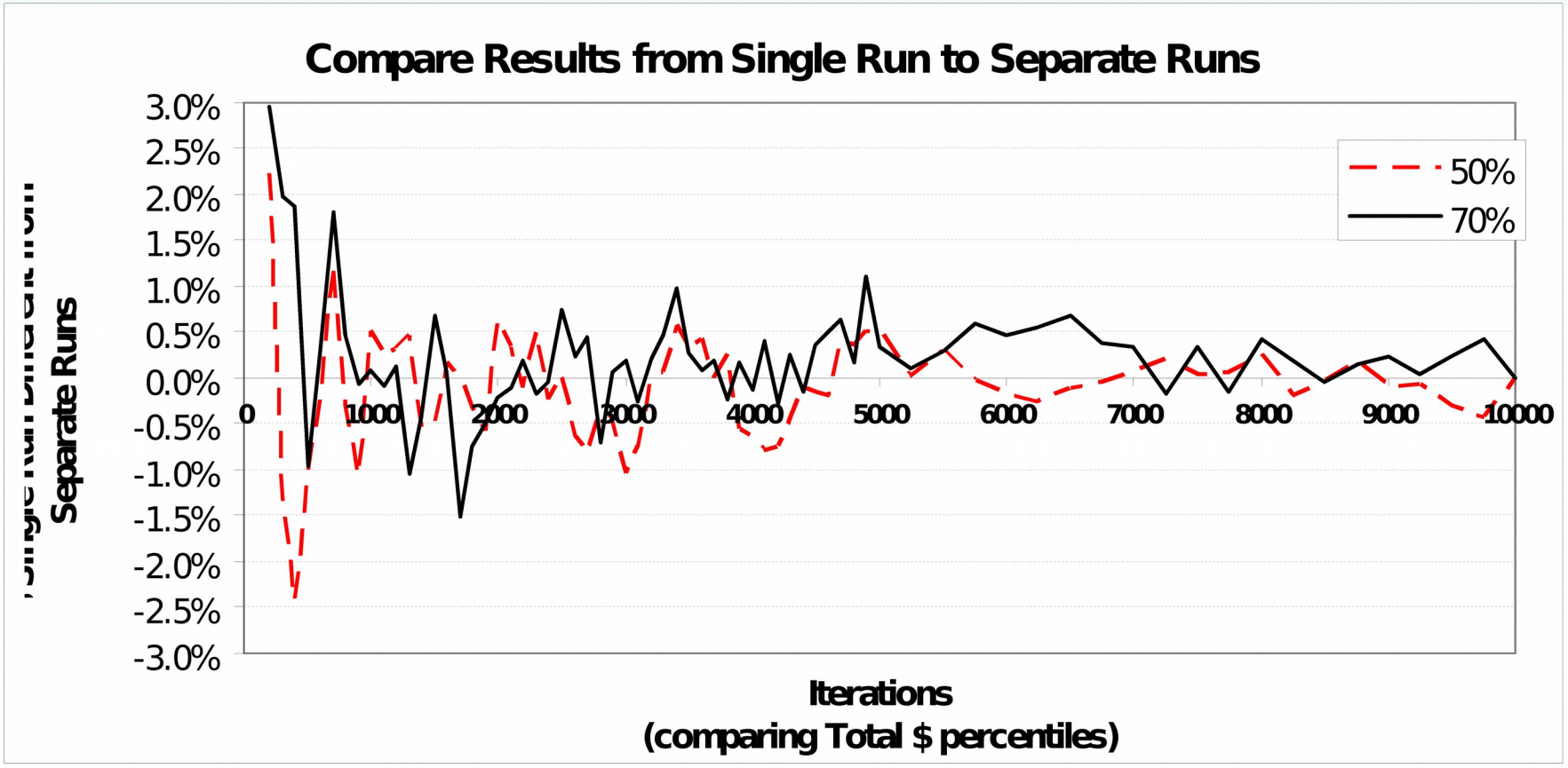
Separate Runs vs A Single Run

- **Option 1**: Generate separate runs:
  - Perform a separate run for each x iterations that will be compared to the 10k run
  - Becomes **extremely time consuming** if any fidelity desired

- **Option 2:** Use data from a single 10k run:
  - Obtain 10k iteration data and calculate statistics based on all 10k
  - Recalculate the statistics based upon the first 200 iterations, first 300, first 400 and so on
    - An alterative is to randomly sample with replacement from the 10k data
  - Does not guarantee distributions are sampled across their entire range
  - Far quicker and easier to manage than Option 1

- **Goal:** Demonstrate that Option 2 is adequate

Is it necessary to perform separate runs when checking for convergence?



- Results from first 200 iterations of a 10k run are compared to an independent run of 200 iterations and so on
- **Conclusion:** analysis of a single 10k run is sufficiently accurate to test for stability

Recap and Way Ahead

- **Recap:**
  - 10k iterations selected as the benchmark
    - Two sources noted
  - Ignore impact of random seed changes
    - Random seed change has a +/- 0.5% impact
  - Use the data from a single 10k simulation run
    - Separate runs more completely sample the distribution, but statistics are generally less than 1% different from statistics calculated from a single 10k run

- **Way Ahead:**
  - Create a tool to calculate the statistics for each sample of interest and compare them to the 10k statistics
  - Design the tool so that the user may “drop in” the iteration data from any source

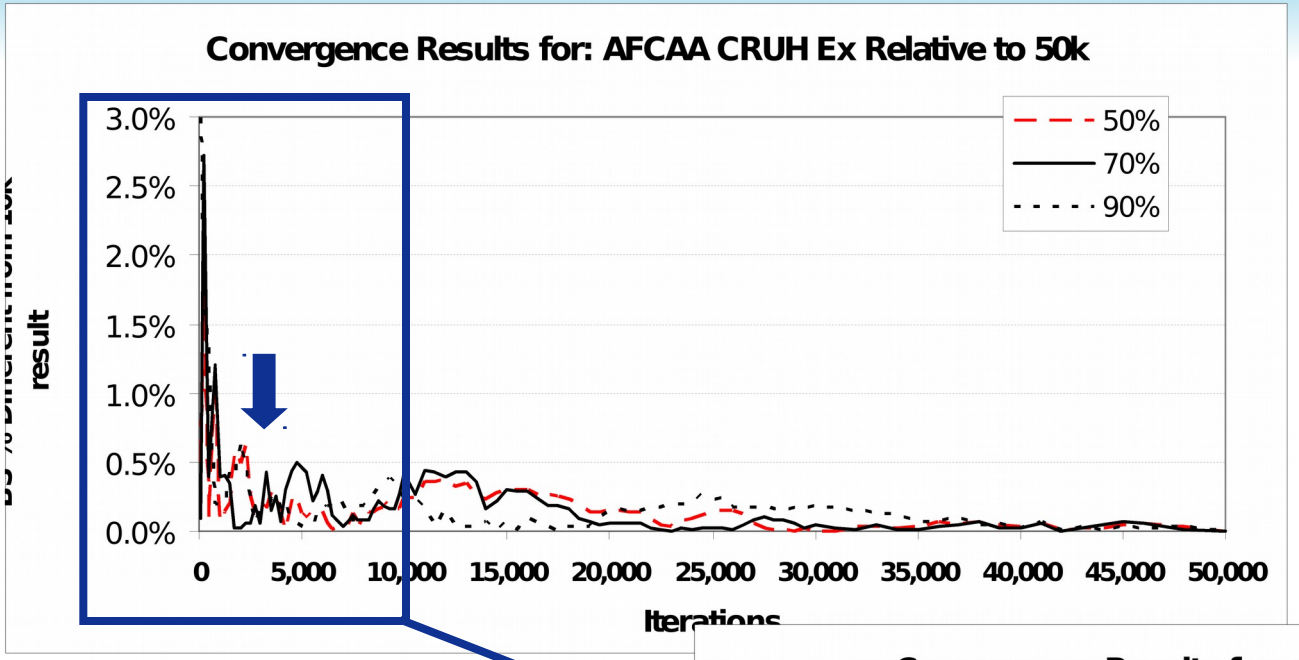
Calculating the Statistics In Excel

	A	B	C	D	E	F	G	H	J	L
49	Iteration	AFCAA CRUH Ex		Iterations	Mean	Stdev	CV	50%	70%	90%
50										
51	1	1,091,662.539		200	933,676.143	226,372.848	0.242	904,350.555	1,044,847.101	1,244,415.362
52	2	727,810.096		300	931,061.0[illegible]9	236,901.474	0.254	893,181.84[illegible]	1,047,547.887	1,250,782.683
53	3	1,205,285.272		400	928,610.71[illegible]	242,462.505	0.261	888,047.982	1,034,861.250	1,262,402.942
54	4	1,111,597.028		500	930,042.410	241,776.774	0.260	889,780.748	1,022,441.553	1,263,843.287
55	5	894,500.526		600	932,321.195	238,122.690	0.255	895,654.455	038,618.409	1,262,550.099

- **Mean = AVERAGE(INDIRECT("B$51:B$" & 50+$D51))**
- **% = LARGE(INDIRECT("B$51:B$" & 50+$D51),ROUND($D51-H$49*$D51,0))**

- **Excel Functions:**
  - **INDIRECT**: allows column D to automatically calculate the statistic from the correct range
  - **LARGE**: finds the value from the correct range for the percentile of interest
- **Copy/paste iteration data from any simulation tool into Column B**
- **Column D can be edited to obtain any granularity of interest**
- **Create additional columns to calculate the % difference from the selected max iterations (in our case, 10k was selected)**
- **Using this approach, the process becomes tool independent**
- **This tool was used to create the charts that follow**

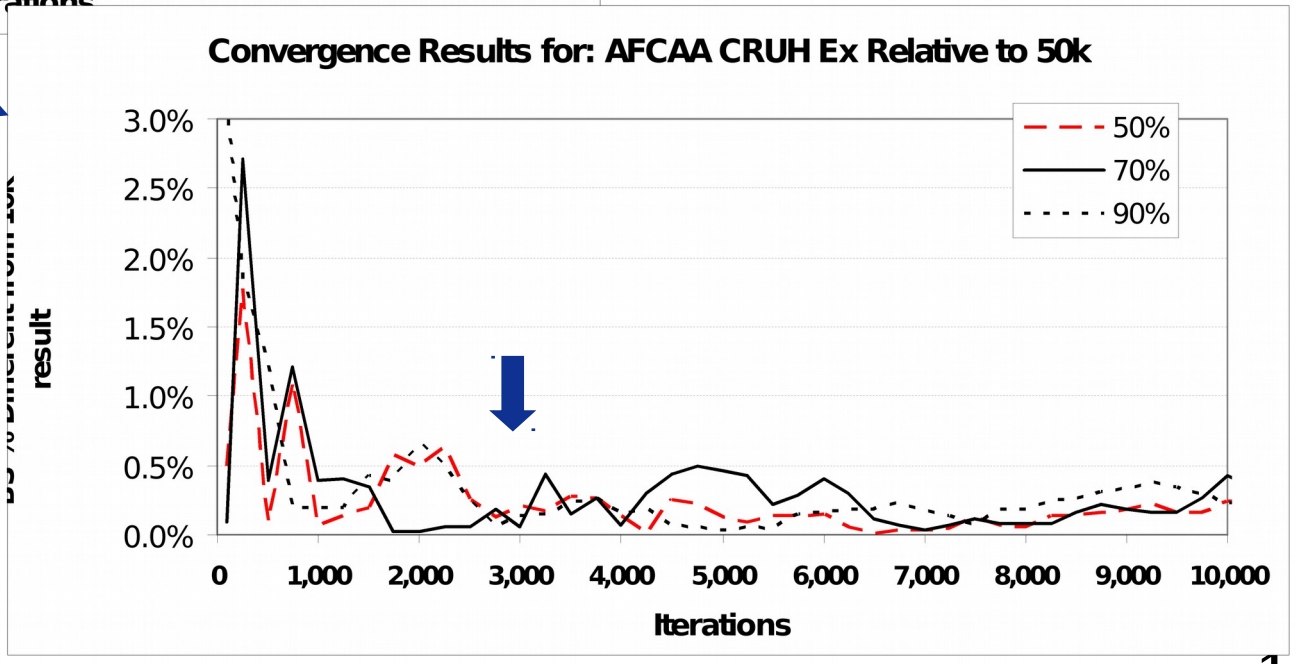
Revisiting The 10k Decision



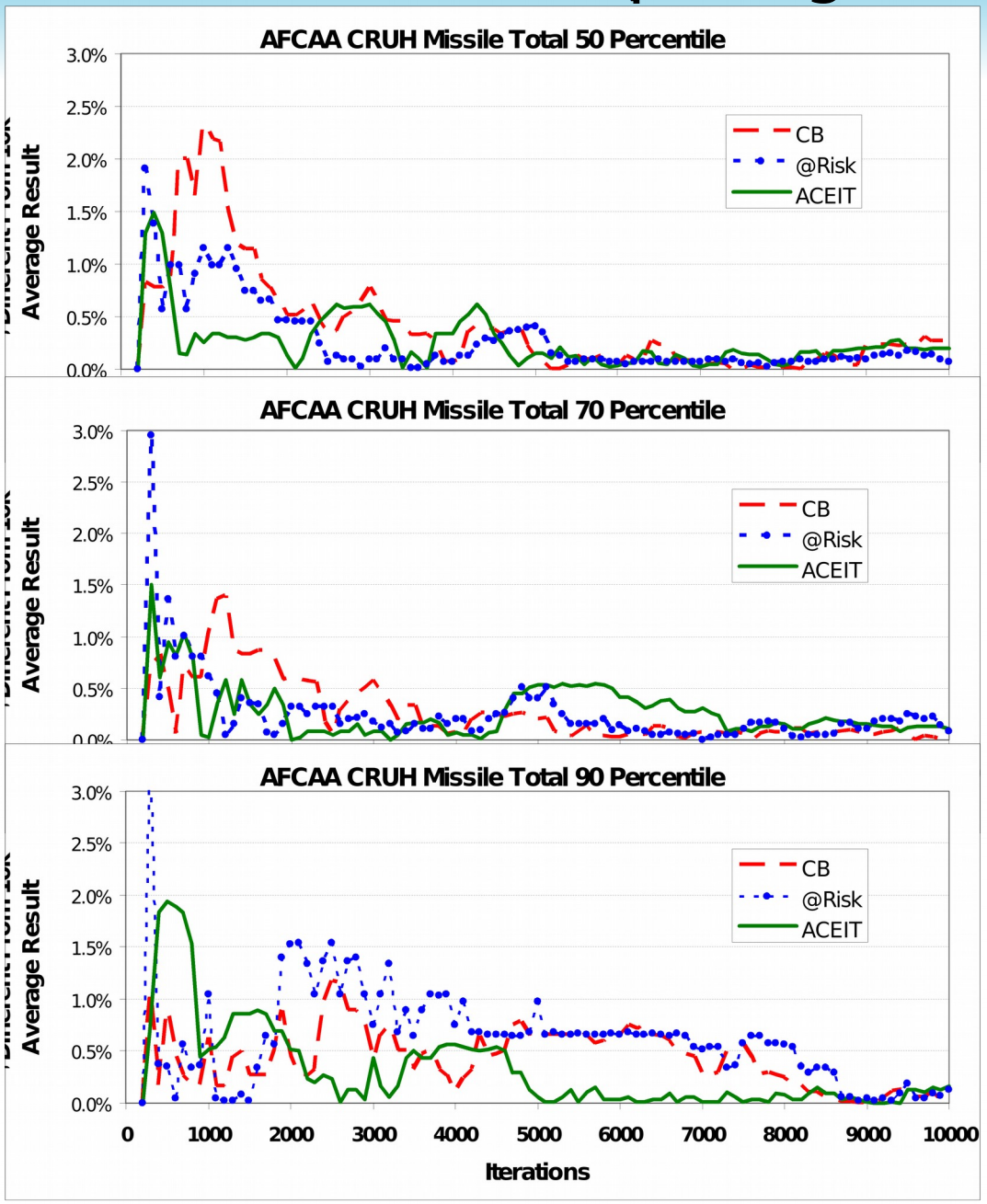
- **Absolute % difference from 50k result is plotted for different confidence levels**
- **Any statistic of interest could be used**

- **If 0.5% different is considered "noise" then anything after 3k qualifies as "accurate**

- **Conclusion:** 10k iterations as the reference for "accurate" stands for this model



Comparing Convergence Across Tools



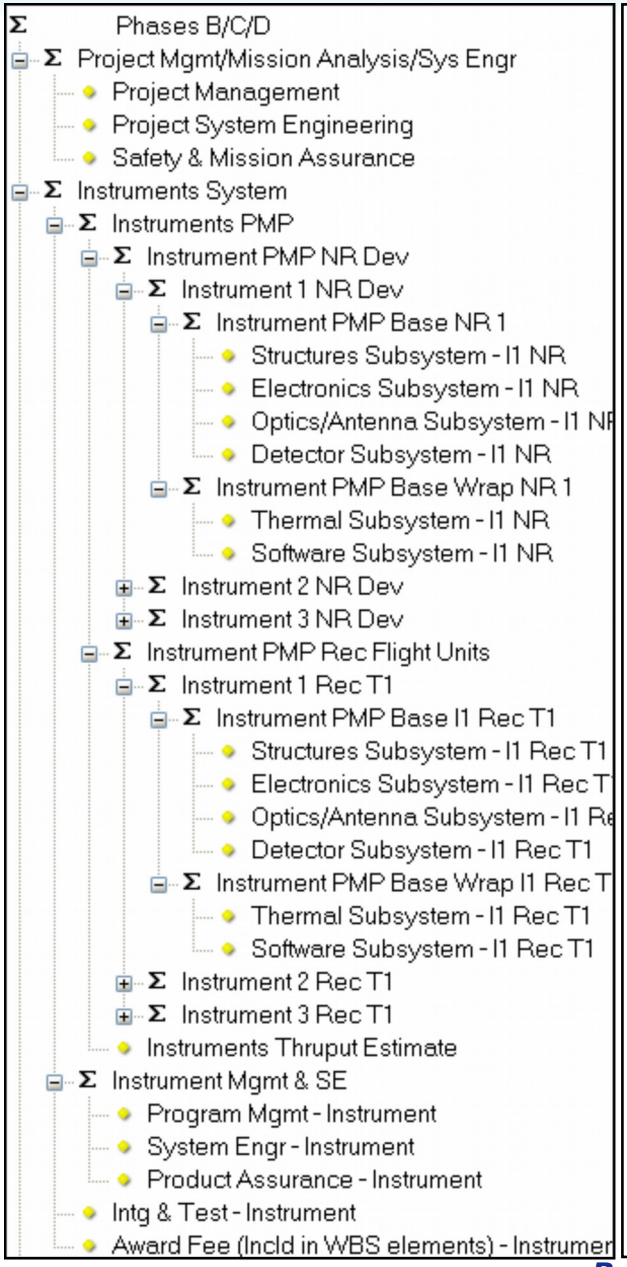
- **Results at each iteration are compared to the average of the three tool results at 10k**
- **Patterns would differ if random seeds changed, but within +/- 0.5%**
- **Conclusion:** All three tools demonstrate similar convergence behavior

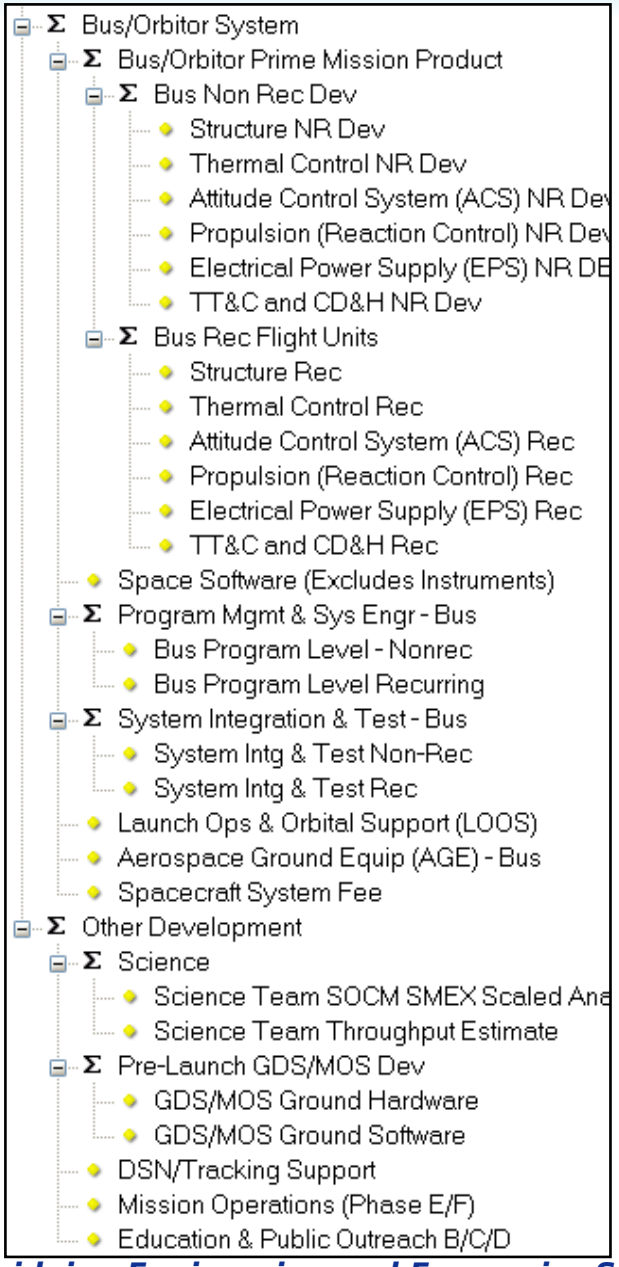
TECOLOTE RESEARCH, INC.

Bridging Engineering and Economics

Convergence For Several Examples

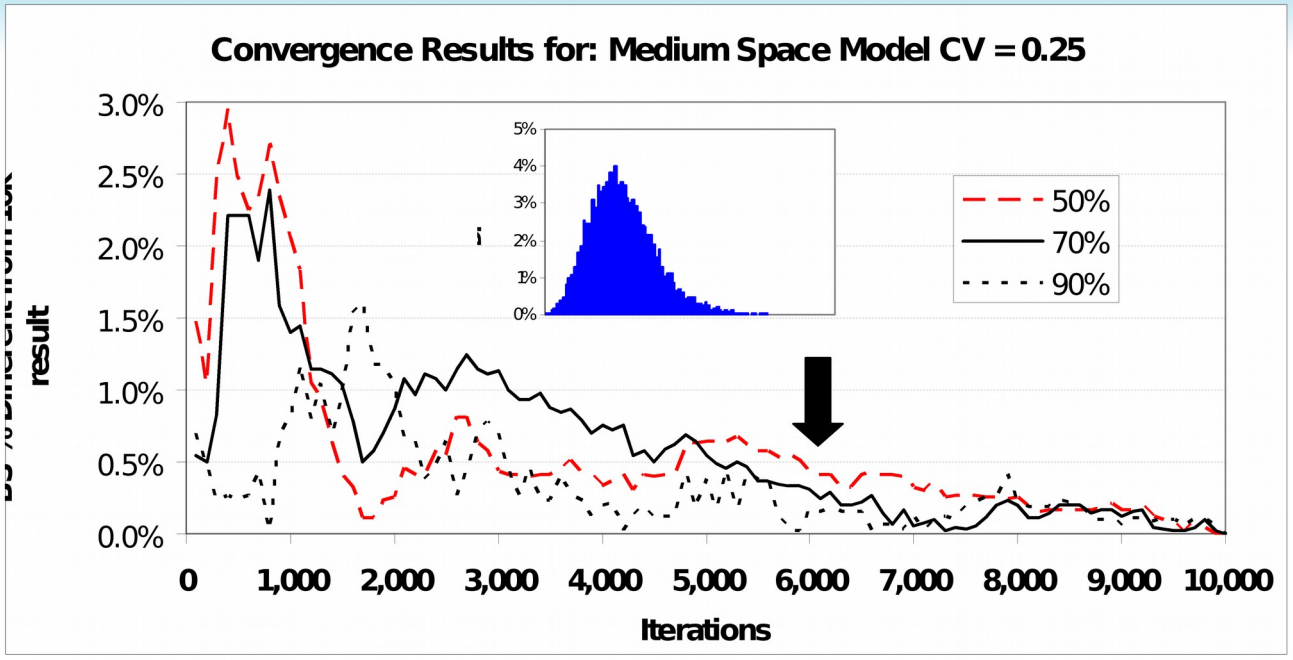
A "Typical" Cost Model





- 73 WBS with estimating methods such as:
  - Non-Linear CERs
  - Linear CERs
  - Factor Relationships
  - Build-up estimates
  - Data from 3rd party tools
  - Throughputs
- Over 150 input variables such as:
  - Labor rates
  - Configuration Inputs (mass, power, etc)
  - Programmatic Inputs (design life, schedule, etc)
  - Factors (overhead wraps, etc)
- Hundreds of intermediate calculations

Identify Iterations For Convergence



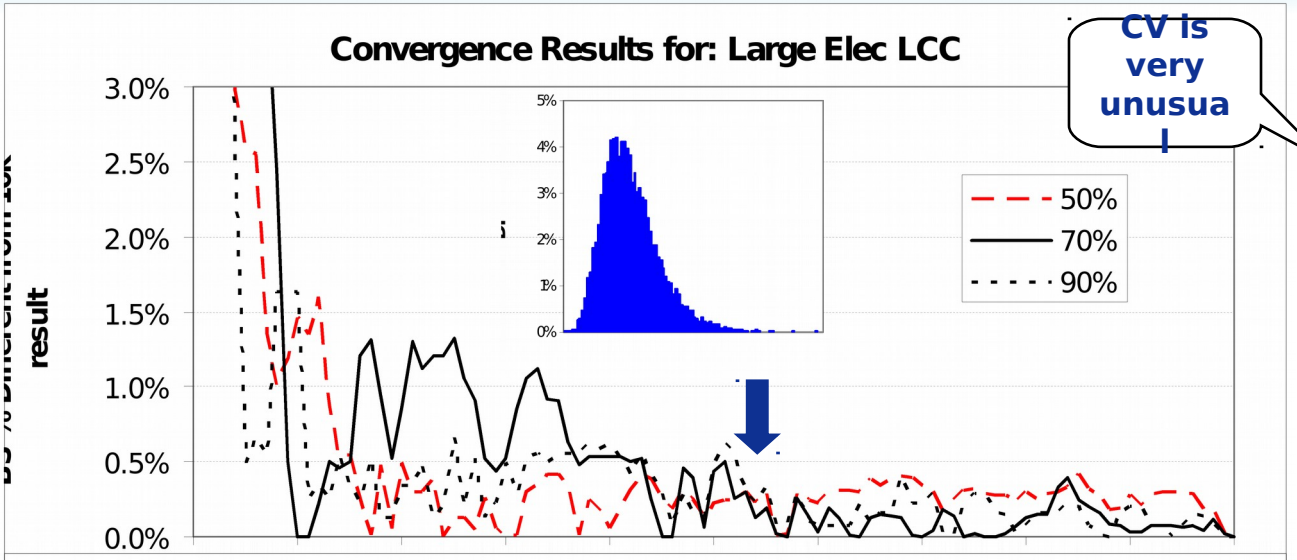
- **6,000 iterations appears sufficient**
- **May be different if anything is changed in the model**
- **Note that distribution shape is not normal**

- **Some metrics to characterize the size of the model:**
  - Elements in the model
  - CV
  - Time to perform 10k iterations
  - Number of distributions
  - Number of applied correlations

Elements in Model	541	Includes WBS, Intermediate calcs and Inputs	
# WBS Elements:	109	Includes Parents	
# WBS Methods	73	67.0%	
10k CV = 0.25	**Min**	**Sec**	**Scale to 10k (Sec)**
10k iterations	1	17.29	77.3
1k Iterations	0	7.73	77.3
500 iterations	0	4.36	87.2
100 iterations		0.94	94.0

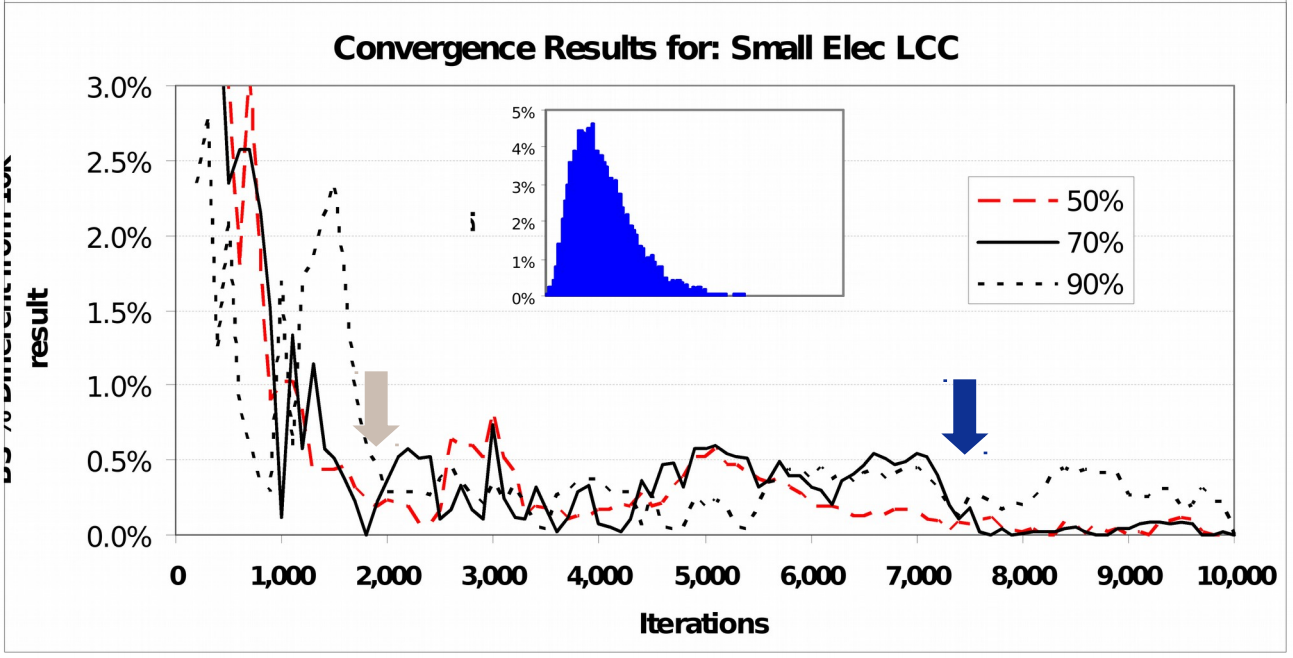
	Distribu-tions	Group Names	Correlation Strength
Count	82	78	78
Unique	3	10	2
Ave # Elements per group		7	

Electronics



Elements in Model	3635	Includes WBS, Intermediate calcs and Inputs	
# WBS Elements:	632	Includes Parents	
# WBS Methods	454	71.8%	
10k CV = 1.25	**Min**	**Sec**	**Scale to 10k (Sec)**
10k iterations	122	11.98	7332.0
1k Iterations	14	13.97	8539.7
500 iterations	6	53.92	8278.4
100 iterations		74.56	7456.0

	Distribu-tions	Group Names	Correlation Strength
Count	291	167	167
Unique	4	2	1
Ave # Elements per group		83	

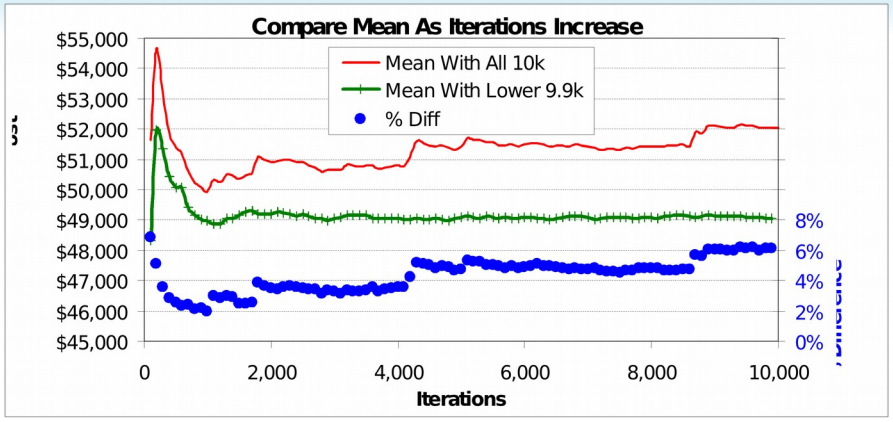


Elements in Model	629	Includes WBS, Intermediate calcs and Inputs	
# WBS Elements:	395	Includes Parents	
# WBS Methods	318	80.5%	
10k CV = 0.38	**Min**	**Sec**	**Scale to 10k (Sec)**
10k iterations	3	52.48	232.5
1k Iterations	0	23.39	233.9
500 iterations	0	11.73	234.6
100 iterations		2.44	244.0

	Distributions	Group Names	Correlation Strength
Count	119	112	112
Unique	3	1	1
Ave # Elements per group		112	

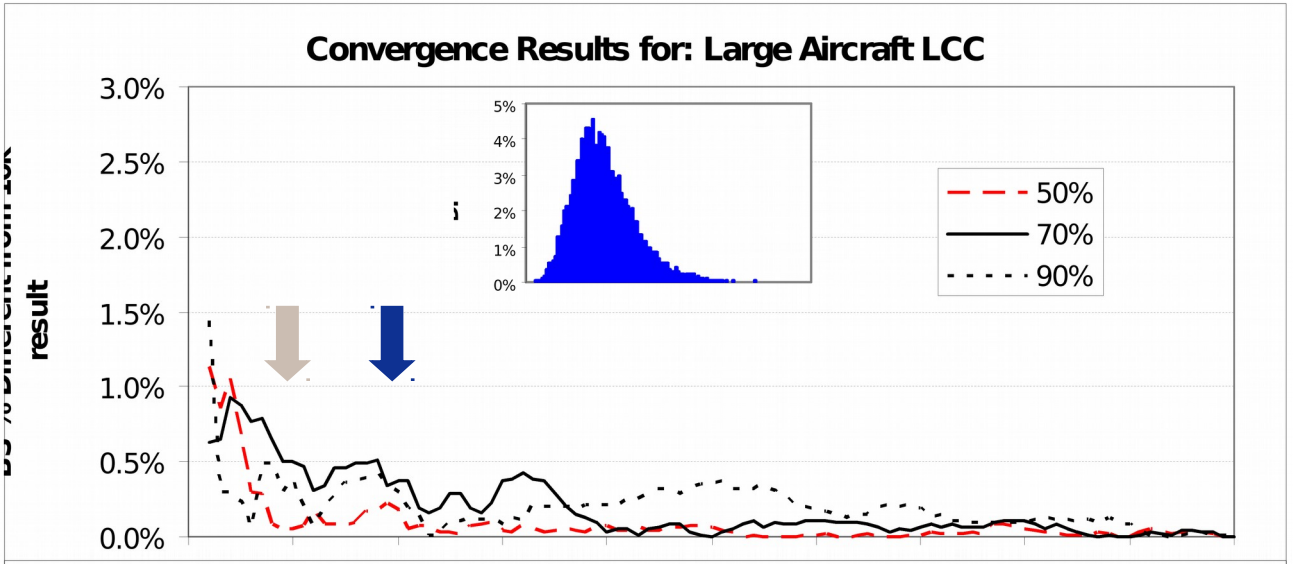
TECOLOTE RESEARCH, INC.

What's Happening in the Large Electronic Simulation?

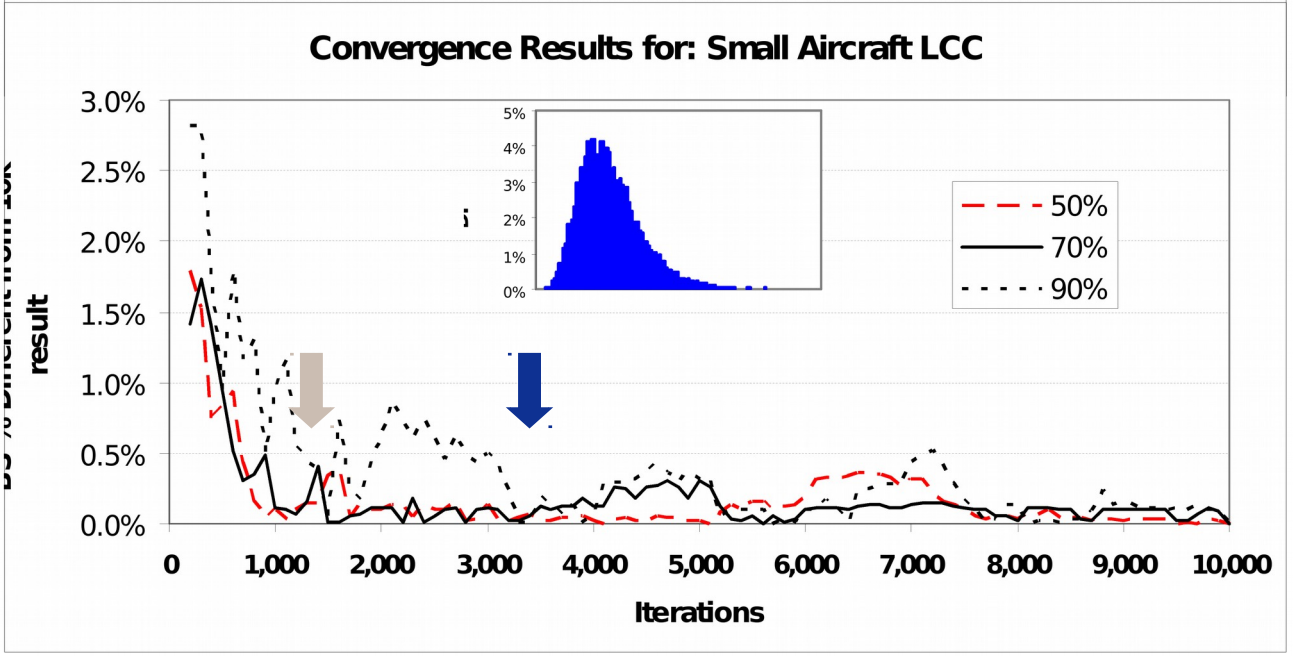


- After 1,000 iterations, the mean climbed (red line) as iterations increased
- CV jumps up dramatically periodically (red line)
- The top 100 results were "stripped" from the simulation and stats

Aircraft LCC

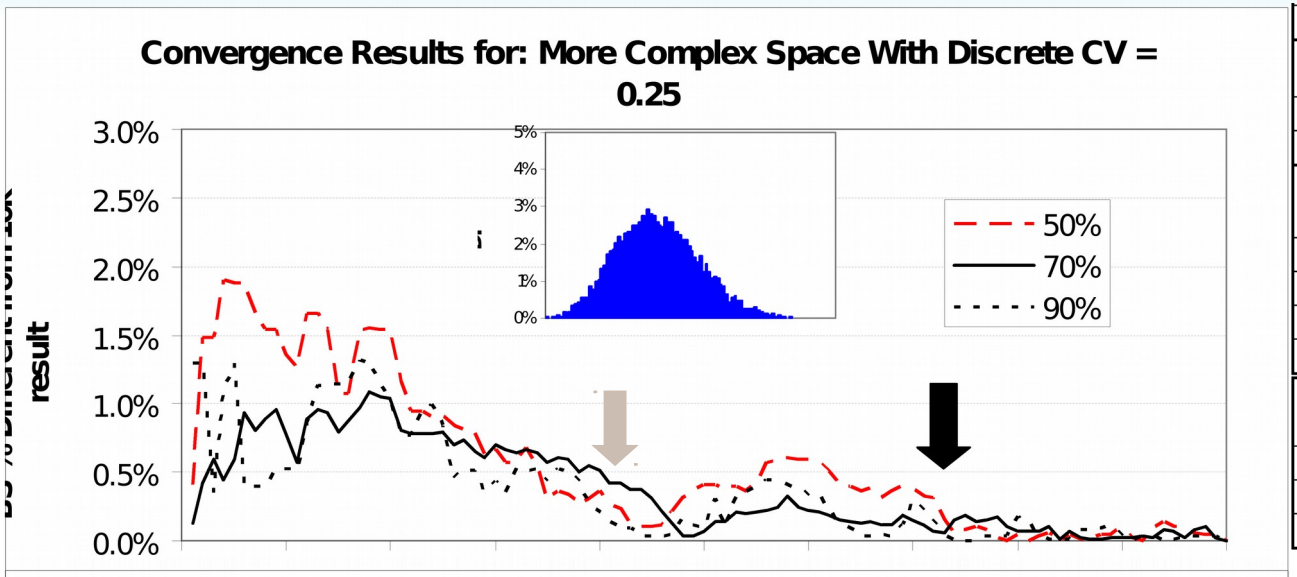


Elements in Model	622	Includes WBS, Intermediate calcs and Inputs	
# WBS Elements:	119	Includes Parents	
# WBS Methods	83	69.7%	
10k CV = 0.09	**Min**	**Sec**	**Scale to 10k (Sec)**
10k iterations	13	26.31	806.3
1k Iterations	1	20.00	800.0
500 iterations	0	40.06	801.2
100 iterations		20.45	2045.0
	Distribu-tions	**Group Names**	**Correlation Strength**
Count	345	144	144
Unique	4	8	4
Ave # Elements per group		18	



Elements in Model	697	Includes WBS, Intermediate calcs and Inputs	
# WBS Elements:	279	Includes Parents	
# WBS Methods	195	69.9%	
10k CV = 0.21	**Min**	**Sec**	**Scale to 10k (Sec)**
10k iterations	11	33.06	693.1
1k Iterations	1	2.77	627.7
500 iterations	0	31.61	632.2
100 iterations		6.45	645.0
	Distribu-tions	**Group Strength**	**Correlation Strength**
Count	129	121	121
Unique	4	4	1
Ave # Elements per group		30	

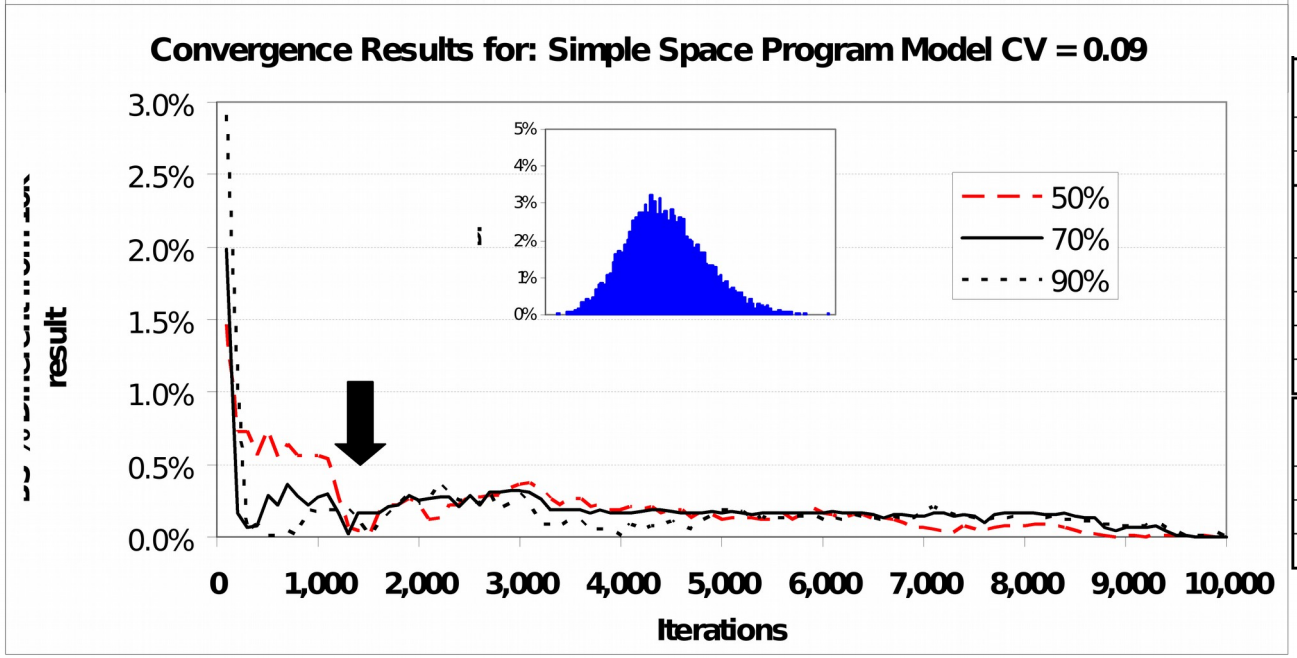
Sample Space Models



Number of:	314	27	8478
Elements in Model	820	Includes WBS, Intermediate calcs and Inputs	
# WBS Elements:	242	Includes Parents	
# WBS Methods	205	84.7%	

10k CV = 0.25	Min	Sec	Scale to 10k (Sec)
10k iterations	39	25.77	2365.8
1k Iterations	3	54.82	2348.2
500 iterations	1	59.46	2389.2
100 iterations		24.77	2477.0

	Distribu-tions	Groups	Group Strength
Count	208	208	208
Unique	1	2	1
Ave # Elements per group		104	

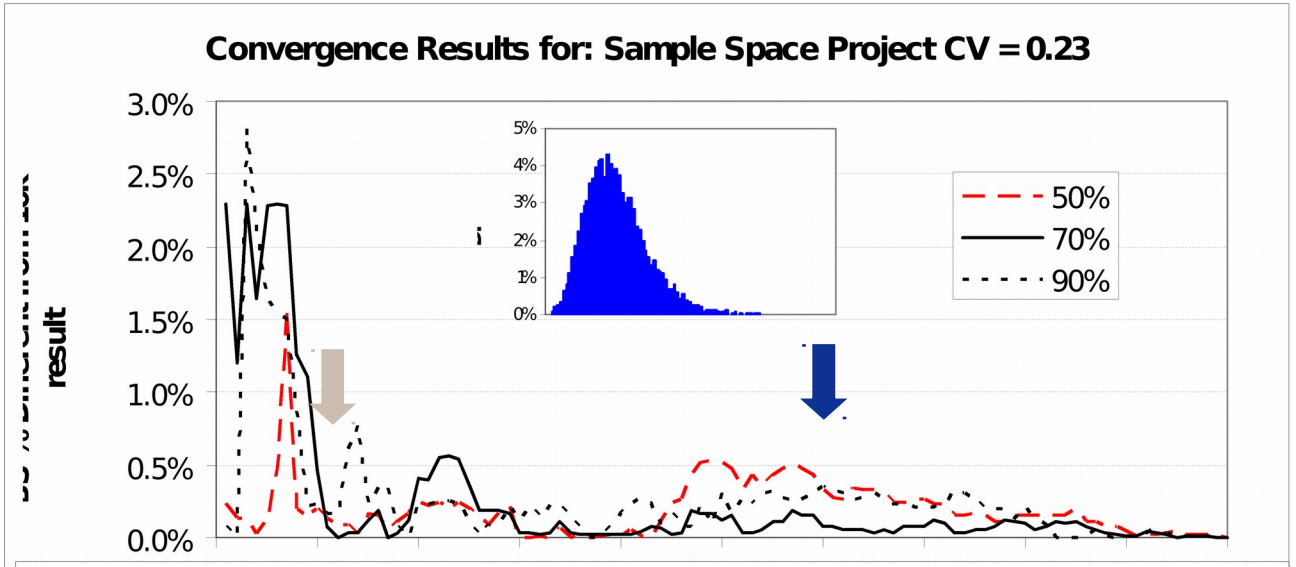


Elements in Model	715	Includes WBS, Intermediate calcs and Inputs	
# WBS Elements:	90	Includes Parents	
# WBS Methods	70	77.8%	

10k CV = 0.09	Min	Sec	Scale to 10k (Sec)
10k iterations	1	24.26	84.3
1k Iterations	0	8.75	87.5
500 iterations	0	4.43	88.6
100 iterations		7.00	700.0

	Distribu-tions	Groups	Group Strength
Count	82	65	65
Unique	2	8	1
Ave # Elements per group		8	

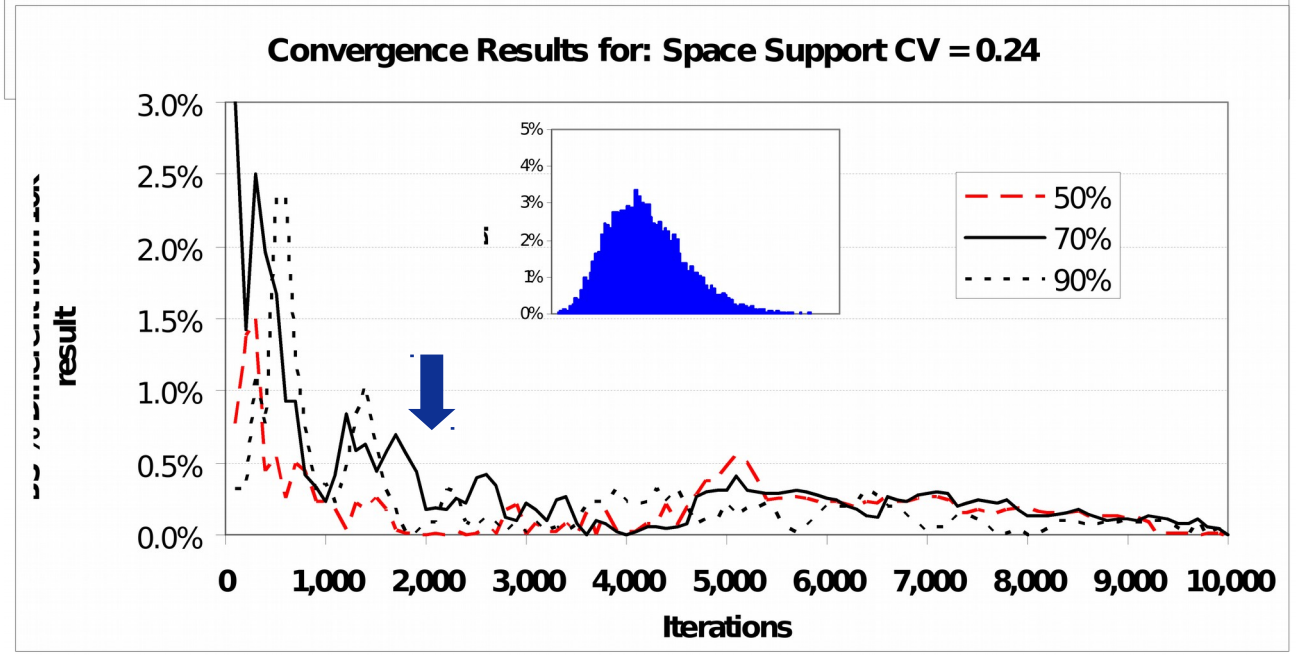
Projects Supporting Space Systems



Elements in Model	433	Includes WBS, Intermediate calcs and Inputs
# WBS Elements:	23	Includes Parents
# WBS Methods	13	56.5%

10k CV = 0.23	Min	Sec	Scale to 10k (Sec)
10k iterations	0	49.27	49.3
1k Iterations	0	5.02	50.2
500 iterations	0	2.56	51.2
100 iterations		0.59	59.0

	Distributions	Groups	Group Strength
Count	30	20	20
Unique	3	3	2
Ave # Elements per group		6	

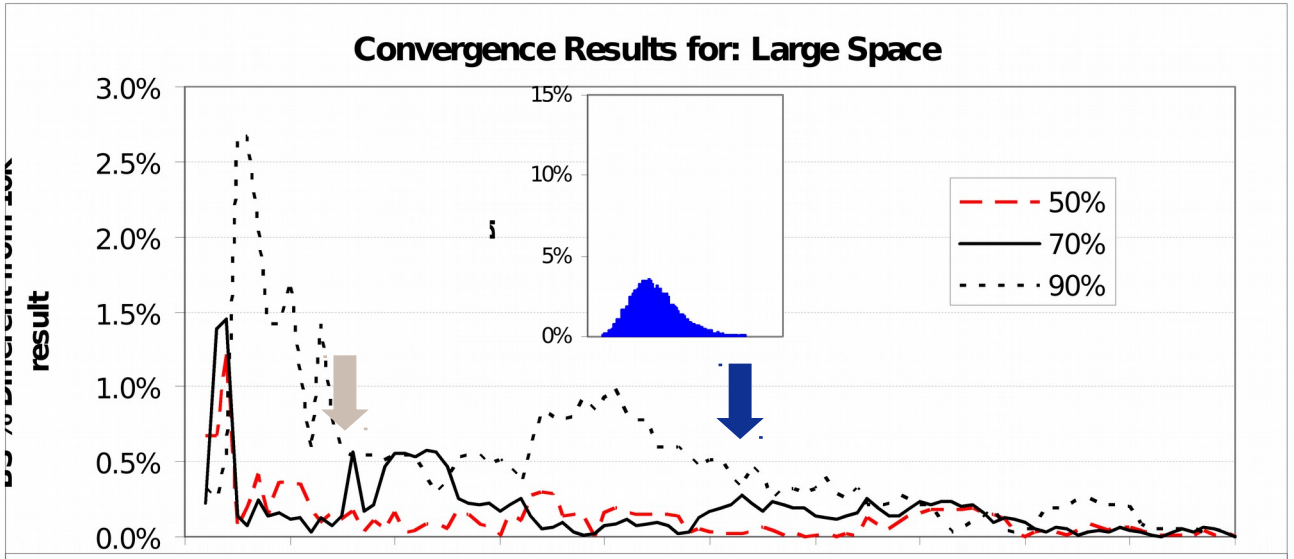


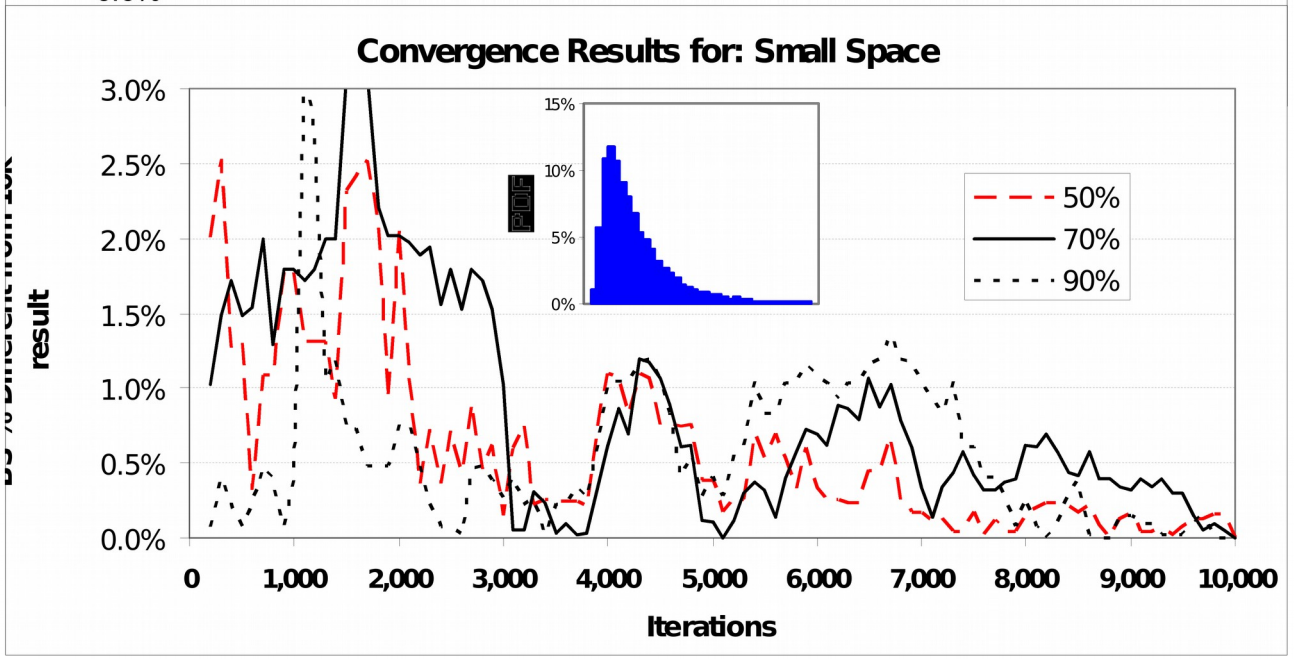
Elements in Model	532	Includes WBS, Intermediate calcs and Inputs
# WBS Elements:	208	Includes Parents
# WBS Methods	186	89.4%

10k CV = 0.24	Min	Sec	Scale to 10k (Sec)
10k iterations	2	37.88	157.9
1k Iterations	0	16.11	161.1
500 iterations	0	8.25	165.0
100 iterations		1.83	183.0

	Distributions	Groups	Group Strength
Count	106	51	51
Unique	1	2	34
Ave # Elements per group		25	

Space Systems

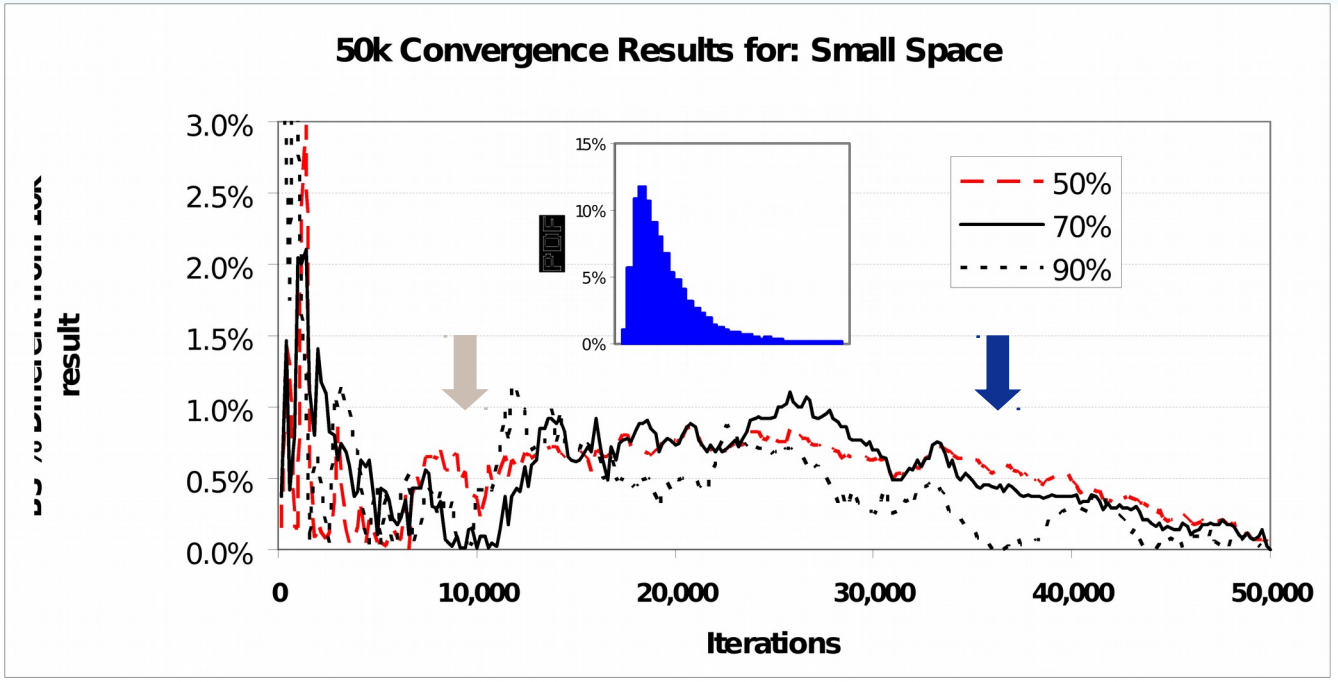




Elements in Model	2214	Includes WBS, Intermediate calcs and Inputs	
# WBS Elements:	957	Includes Parents	
# WBS Methods	732	76.5%	
10k CV = 0.2	**Min**	**Sec**	**Scale to 10k (Sec)**
10k iterations	36	54.08	2214.1
1k Iterations	8	1.88	4818.8
500 iterations	1	18.95	1579.0
100 iterations		16.28	1628.0
	Distribu-tions	**Group Names**	**Correlation Strength**
Count	486	397	396
Unique	3	2	7
Ave # Elements per group		198	

Elements in Model	82	Includes WBS, Intermediate calcs and Inputs	
# WBS Elements:	9	Includes Parents	
# WBS Methods	7	77.8%	
10k CV = 0.72	**Min**	**Sec**	**Scale to 10k (Sec)**
10k iterations	0	10.44	10.4
1k Iterations	0	1.11	11.1
500 iterations	0	0.58	11.6
100 iterations		0.23	23.0
	Distribu-tions	**Group Names**	**Correlation Strength**
Count	14	3	3
Unique	3	1	1
Ave # Elements per group		3	

Why Does the Small Space Model Require So Many Iterations?



- **Model based upon following equation:**
  - 0.6636*V1^0.6567 * V2^0.1555 * V3^0.03226 * V4^0.4409 * V5^0.9142 * V6^-0.2879
  - Uncertainty on each variable
- **CER result used to estimate other cost elements using uncertain factor relationships**

One of the smallest models, takes the most iterations

Concluding Comments

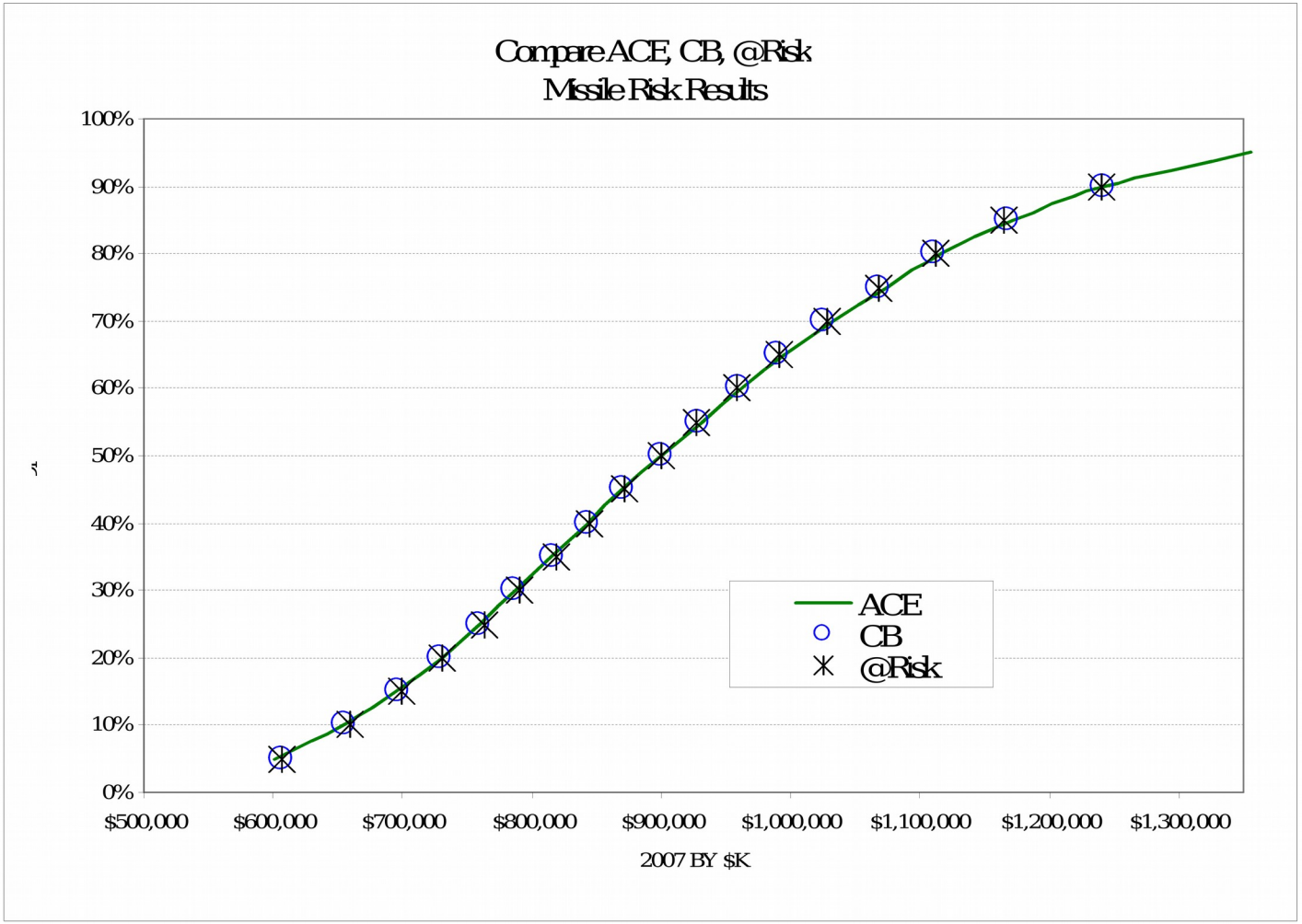
- **Convergence was defined as the number of iterations required such that statistic of interest stays within 0.5% of the 10k result**
  - 50, 70, 90 percentile selected in this study as basis for testing for convergence
- **Simple Excel tool provides a consistent, tool independent way to test for convergence**
- **Observations:**
  - None of the models generated a Normal distribution at the total level
  - Can ignore impact of random seed changes
  - Convergence can be estimated from a single 10k simulation run
  - Models tested converged faster than analytic formula suggests, possibly due using Latin Hypercube over Monte Carlo
  - Contrary to the analytic approach, more iterations are required as percentile increases
  - CV more important than # of elements in model when assessing iteration requirement
    - 10k iterations may be insufficient if model CV is high, i.e. > 0.6
- **How many iterations are required?**
  - Unfortunately, the answer is: it depends
  - Use a simple, consistent method to find out

TECOLOTE RESEARCH, INC.

Bridging Engineering and Economics

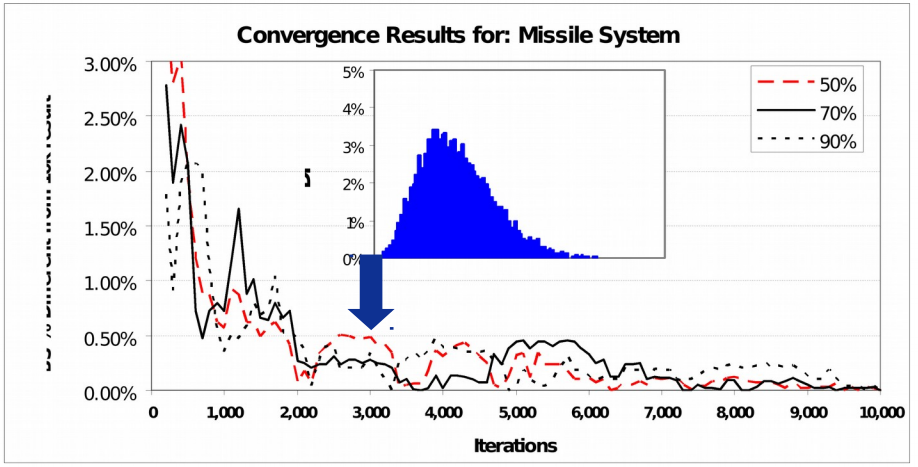
Backup

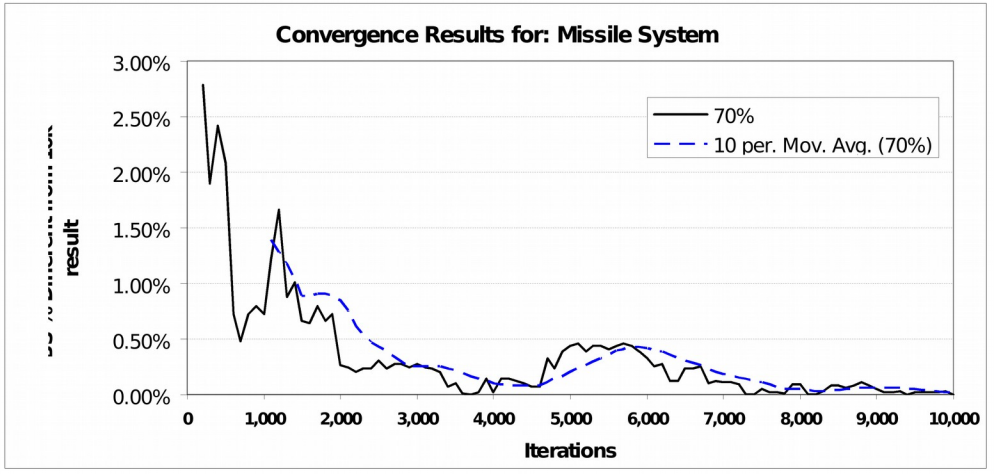
AFCAA CRUH 10k Iteration Example Results

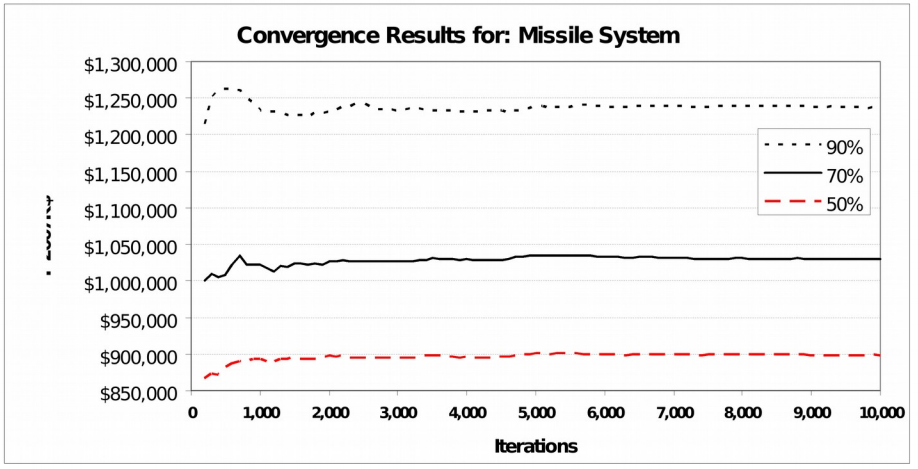


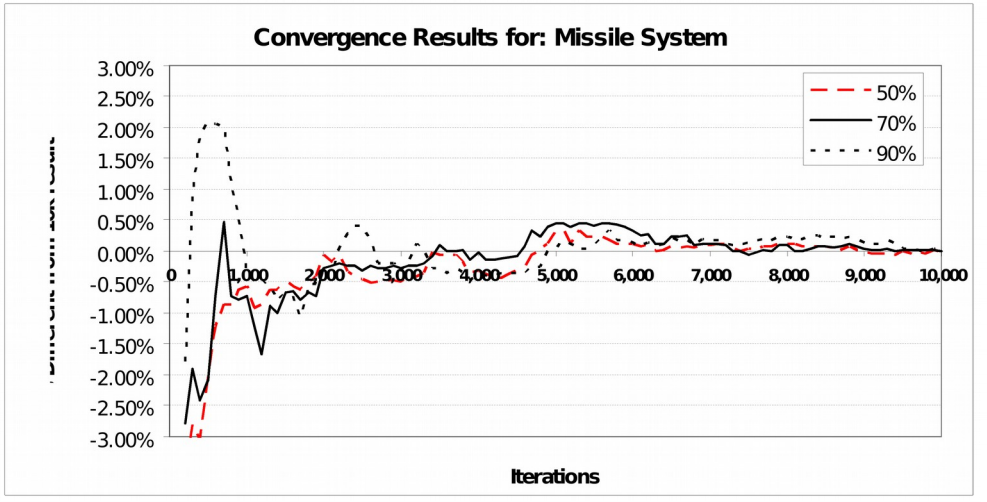
- **Results are tool independent**
- **The handbook does not endorse or recommended any specific tool**

Different Ways to Present Results



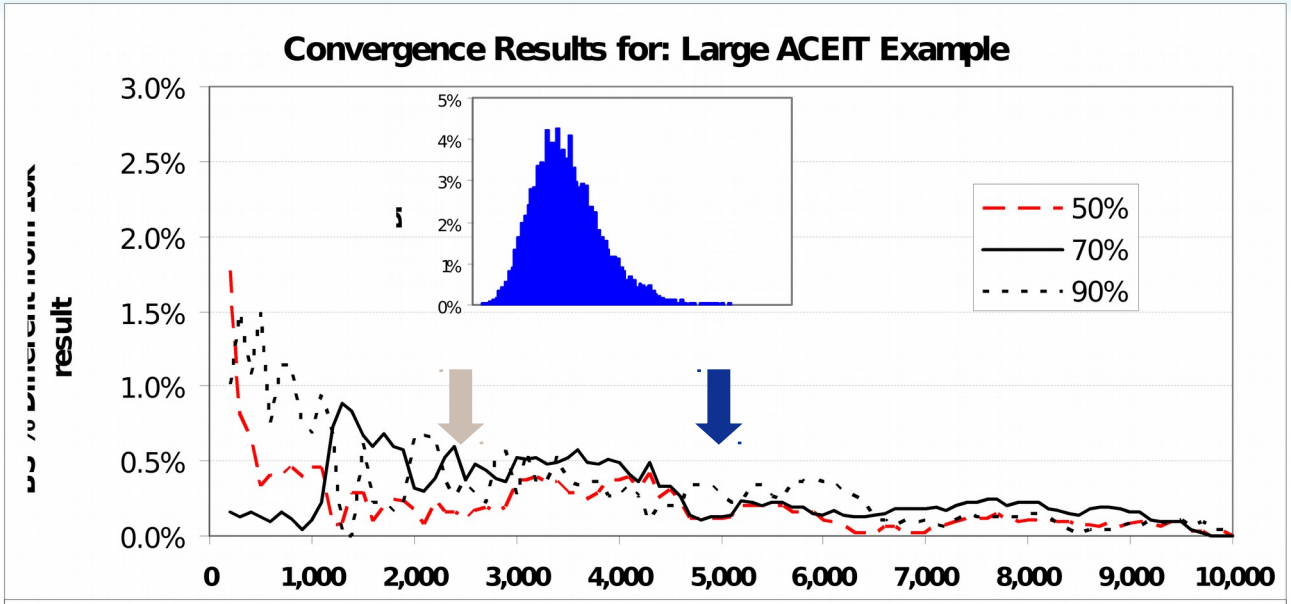






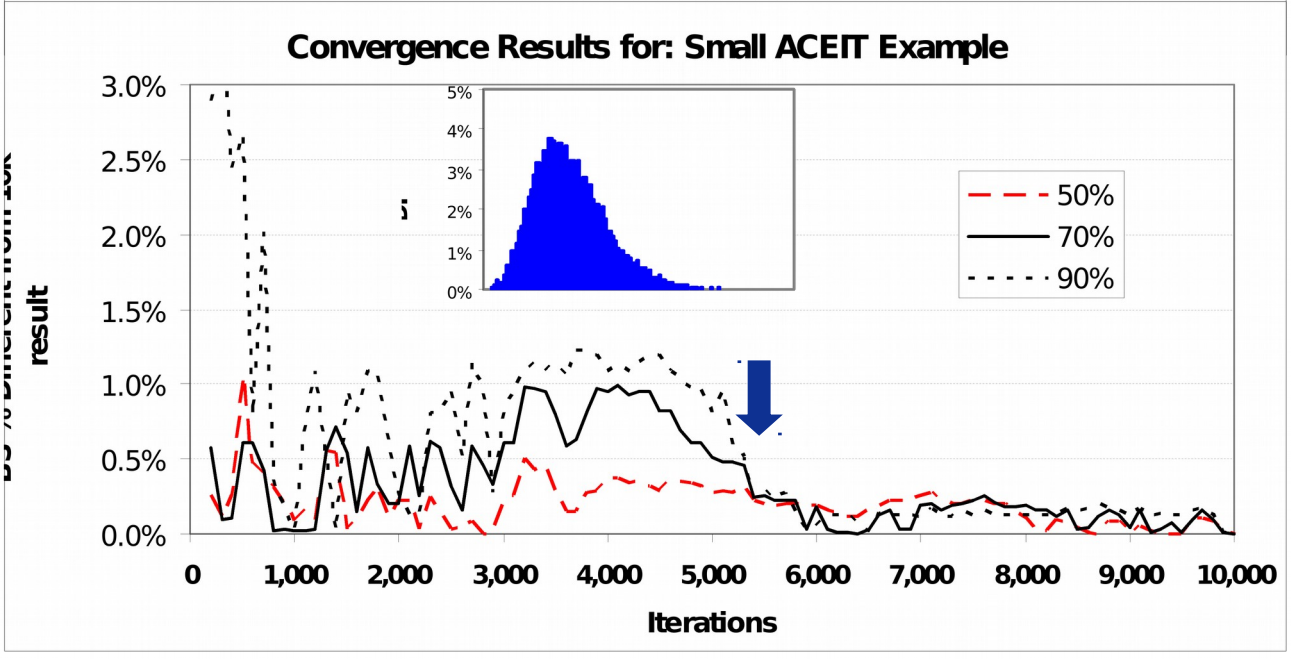
- **Derived from evaluating the iteration data from a 10k run**
- **Appears that for this model (AFCAA CRUH Ex), 2-3 k iterations are sufficient**
- **Conclusion:** Upper left selected as standard way to present analysis

ACEIT Example Files



Elements in Model	313	Includes WBS, Intermediate calcs and Inputs	
# WBS Elements:	91	Includes Parents	
# WBS Methods	60	65.9%	
10k CV = 0.19	**Min**	**Sec**	**Scale to 10k (Sec)**
10k iterations	1	21.30	81.3
1k Iterations	0	8.17	81.7
500 iterations	0	4.19	83.8
100 iterations		0.97	97.0

	Distribu-tions	Group Names	Correlation Strength
Count	75	48	48
Unique	7	10	9
Ave # Elements per group		4	



Elements in Model	33	Includes WBS, Intermediate calcs and Inputs	
# WBS Elements:	6	Includes Parents	
# WBS Methods	4	66.7%	
10k CV = 0.27	**Min**	**Sec**	**Scale to 10k (Sec)**
10k iterations	0	1.02	1.0
1k Iterations	0	0.14	1.4
500 iterations	0	0.09	1.8
100 iterations		0.05	5.0

	Distribu-tions	Group Names	Correlation Strength
Count	6	2	2
Unique	3	1	2
Ave # Elements per group		2	